Crossing-the-Bridge Noodles: Thoughts on the Gap Between Data and Metadata

Kevin D. Keck
Lawrence Berkeley National Laboratory

SC32 Plenary, Kunming
May 24, 2010
Why Does It Matter?
Status Quo

“Cylinders of Excellence”
Solutions

- SQL-MM MDR (ISO 13249-8)
  - a great step
  - still a silo

- Web-addressability scheme for Databases
  - standard URI scheme for referring to DBs (and tables, columns, types, etc. therein)
  - could solve one direction

- Tackle federation
  - integrate the silos
Web-addressability

- There are many URL schemes for RDBMSs – odbc:, jdbc:, dbi:, file:
- There is no mechanism to determine whether two such URLs point to the same DB
- There is no standard programmatic mechanism to obtain a URL for access for a given client
- Would be analogous to DOIs, ISBNs, etc.
- Would ideally be supported directly by DBMSs following ISO SQL standard
Complications

- Will SQL-MM MDR work in most commonly-used RDBMS systems?
- What about the cloud DBs?
  - Google File System (GFS)
  - Hadoop Distributed File System (HDFS)
- What about streaming data?
  - Sensorpedia
Contact Information

XMDR Project
xmdr.org

Kevin D. Keck
kdkeck@lbl.gov